What is claimed is:

1. A video encoding apparatus comprising:

an orthogonal transform circuit for orthogonally transforming an input picture signal to obtain a plurality of transform coefficients;

a first local decoder for outputting first transform coefficients for a fine motion compensation prediction picture on the basis of a previous picture;

a second local decoder for outputting second transform coefficients for a coarse motion compensation prediction picture on the basis of a current picture corresponding to the input picture signal;

means for detecting a degree of motion compensation prediction in said second local decoder;

a selector for selectively outputting the first and second transform coefficients in accordance with the degree of motion compensation prediction;

a first calculator for calculating a difference between the transform coefficients of said orthogonal transform circuit and ones of the first and second transform coefficients which are selected by said selector, and outputting a motion compensation prediction error signal;

a first quantizer for quantizing the motion compensation prediction error signal from said first adder and outputting a first quantized motion compensation prediction error signal;

a second calculator for calculating a difference between the second transform coefficients from said second local decoder and the transform coefficients from the orthogonal transform circuit, and outputting a second motion compensation prediction error signal;

a second quantizer for quantizing the motion compensation prediction error signal from said second calculator, and outputting a second quantized motion compensation prediction error signal; and

an encoder for encoding the first and second quantized motion compensation prediction error signals and outputting encoded signals.

2. An apparatus according to claim 1, wherein said orthogonal transform circuit divides the input picture signal into a plurality of blocks each having a size of N × N pixels, orthogonally transforms the picture signal in units of blocks each having the size of N × N pixels, and outputs N × N transform coefficients to said first and second adders in units of blocks.

3. An apparatus according to claim 1, wherein said means for detecting degree of motion compensation prediction compares the second transform coefficient from said second local decoder with the transform coefficient from said orthogonal transform circuit and outputs a detection signal corresponding to the degree of motion compensation prediction, and said selector

selects the first transform coefficient from said first local decoder in response to a detection signal representing that motion compensation prediction is correct, and selects the second transform coefficient from said second local decoder in response to a detection signal representing that motion compensation prediction is incorrect.

4. An apparatus according to claim 1, which comprises a motion vector detector for detecting a motion vector from the input picture signal, said first local decoder comprises a first dequantizer for dequantizing the first quantized motion compensation prediction error signal from said first quantizer and outputting a first output signal, a first adder for adding the first output signal from the first dequantizer and the transform coefficient selected by said selector and outputting a first local decoded signal, and a first motion compensation prediction circuit for obtaining the first transform coefficient on the basis of the first local decoded signal from said first adder and the motion vector, and said second local decoder comprises a second dequantizer for dequantizing the second quantized motion compensation prediction error signal from said second quantizer, and outputting a second output signal, a second adder for adding the second output signal from said second dequantizer and the second transform coefficient and

outputting a second local decoded signal, and a second motion compensation prediction circuit for obtaining the second transform coefficient on the basis of the second local decoded signal from said second adder and the motion vector.

5. A video encoding apparatus comprising:

an orthogonal transform circuit for dividing an input video signal into a plurality of blocks each containing N × N pixels and orthogonally transforming the input video signal in units of blocks to obtain a plurality of transform coefficients divided in spacial frequency bands;

a first motion prediction processing section for performing motion compensation prediction processing for the plurality of transform coefficients in order to obtain an upper-layer motion compensation prediction signal having the number of data enough to obtain a high image quality;

a second motion prediction processing section for performing motion compensation prediction processing for the plurality of transform coefficients in order to obtain a lower-layer motion compensation prediction signal upon reducing the number of data;

a decision section for deciding in motion compensation on the basis of the lower-layer motion compensation prediction signal whether motion compensation prediction is correct;

a selector for selecting the upper-layer motion compensation prediction signal in response to a decision representing a correct motion compensation prediction from said decision section, and the lower-layer motion compensation prediction signal in response to a decision representing an incorrect motion compensation prediction; and

an encoder for encoding one of the upper-layer motion compensation prediction signal and the lower-layer motion compensation prediction signal which is selected by said selector.

6. A video encoding apparatus for realizing SNR scalability in M layers, comprising:

an orthogonal transform circuit for dividing an input video signal into a plurality of blocks each containing N × N pixels and orthogonally transforming the input video signal in units of blocks to obtain a plurality of transform coefficients divided in spacial frequency bands;

a first motion compensation prediction processing section for performing motion compensation prediction processing for the plurality of transform coefficients in order to obtain an mth-layer (m = 2 to M) motion compensation prediction signal;

a second motion compensation prediction processing section for performing motion compensation prediction processing for the plurality of transform coefficients

in order to obtain an (m-1)th-layer motion compensation prediction signal;

switching means for selecting the mth-layer motion compensation prediction signal of said first motion compensation prediction processing section in order to obtain an mth-layer prediction value when a quantized output from said second motion compensation prediction processing section is 0, and switching between the mth-layer motion compensation prediction signal and the (m-1)th-layer motion compensation prediction signal in units of transform coefficients in order to select the (m-1)th-layer motion compensation prediction signal when the quantized output is not less than 1;

means for calculating a difference signal between an (m-1)th-layer dequantized output from said second motion compensation prediction processing section and an mth-layer motion compensation prediction error signal obtained by a difference between the mth-layer motion compensation prediction signal and the transform coefficient from said orthogonal transform circuit; and

encoding means for quantizing and encoding the difference signal to output an encoded bit stream.

7. An apparatus according to claim 6, wherein said switching means comprises means for switching an mth-layer (m = 2 to M) quantization matrix between a transform coefficient corresponding to a quantized output of zero in the (m-1)th layer and a transform

coefficient corresponding to a quantized output of not less than 1.

8. An apparatus according to claim 6, further comprising means for, when quantized values of two-dimensional orthogonal transform coefficients in the mth layer (m = 2 to M) are to be arranged into a one-dimensional series, arranging a transform coefficient corresponding to a quantized output of not less than 1 in the (m-1)th layer and then a transform coefficient corresponding to a quantized output of 0 in the (m-1)th layer in an order named.

9. An apparatus according to claim 6, wherein said orthogonal transform circuit comprises means for transforming an arbitrary-shape picture into a transform coefficient in accordance with an alpha-map signal for discriminating a background of an input picture from an object thereof, said encoding means comprises means for encoding the alpha-map signal, said first motion compensation prediction processing section comprises means for inversely transforming the transform coefficient in accordance with the encoded alpha-map signal to reconstruct the arbitrary-shape picture.

10. A video encoding/decoding system comprising:

a video encoding apparatus for realizing SNR (Signal to Noise Ratio) scalability in M layers, which includes,

an orthogonal transform circuit for dividing an input video signal into a plurality of blocks each containing N × N pixels and orthogonally transforming the input video signal in units of blocks to obtain a plurality of transform coefficients divided in spacial frequency bands,

a first motion compensation prediction processing section for performing motion compensation prediction processing for the plurality of transform coefficients in order to obtain an mth-layer (m = 2 to M) motion compensation prediction signal,

a second motion compensation prediction processing section for performing motion compensation prediction processing for the plurality of transform coefficients in order to obtain an (m-1)th-layer motion compensation prediction signal,

switching means for selecting the mth-layer motion compensation prediction signal of said first motion compensation prediction processing section in order to obtain an mth-layer prediction value when a quantized output from said second motion compensation prediction processing section is 0, and switching between the mth-layer motion compensation prediction signal and the (m-1)th-layer motion compensation prediction signal in units of transform coefficients in order to select the (m-1)th-layer motion compensation prediction signal when the quantized output is not less

than 1,

means for calculating a difference signal between an (m-1)th-layer dequantized output from said second motion compensation prediction processing section and an mth-layer motion compensation prediction error signal obtained by a difference between the mth-layer motion compensation prediction signal and the transform coefficient from said orthogonal transform circuit, and

encoding means for quantizing and encoding the difference signal to output an encoded bit stream; and

a video decoding apparatus which includes,

means for extracting codes up to a code in the mth (m = 2 to M) layer from the encoded bit stream from said video encoding apparatus,

decoding means for decoding the codes of respective layers up to the mth layer,

dequantization means for dequantizing, in the respective layers, the quantized values decoded by said decoding means,

switching means for switching the mth-layer (m = 2 to M) motion compensation prediction value and the (m-1)th-layer motion compensation prediction value in units of transform coefficients, and outputting the mth-layer motion compensation prediction value for the quantized output of 0 in the (m-1)th layer and the

(m-1)th-layer motion compensation prediction value for the quantized output of not less than 1 in the (m-1)th layer in units of transform coefficients in order to obtain the mth-layer prediction value, and

means for adding the mth-layer motion compensation prediction value and the (m-1)th-layer motion compensation prediction value to reconstruct the mth-layer motion compensation prediction error signal.

11. A system according to claim 10, further comprising switching means for switching an mth-layer (m = 2 to M) quantization matrix between a transform coefficient corresponding to a quantized output of zero in the (m-1)th layer and a transform coefficient corresponding to a quantized output of not less than 1.

12. A system according to claim 10, wherein said decoding means comprises means for decoding an alpha-map signal for discriminating a background of an input picture from an object thereof, and means for inversely transforming the arbitrary-shape picture transformed into the transform coefficient in accordance with the encoded alpha-map signal to reconstruct the arbitrary-shape picture.

13. A video encoding apparatus comprising:

an orthogonal transform circuit for dividing an input video signal into a plurality of blocks each containing N × N pixels and orthogonally transforming an arbitrary-shape picture in units of blocks in

accordance with an alpha-map signal for discriminating a background of an input picture from at least one object thereof to obtain a plurality of transform coefficients;

a first motion compensation prediction processing section having frame memories respectively corresponding to the background and at least one object, means for inversely transforming the transform coefficient in accordance with the alpha-map signal to reconstruct the arbitrary-shape picture, and means for performing motion compensation prediction processing for the plurality of transform coefficients, said first motion compensation prediction processing section being adapted to obtain an upper-layer motion compensation prediction signal having the number of data enough to obtain a high image quality;

a second motion compensation prediction processing section having frame memories respectively corresponding to the background and at least one object, means for inversely transforming the transform coefficient in accordance with the alpha-map signal to reconstruct the arbitrary-shape picture, and means for performing motion compensation prediction processing for the plurality of transform coefficients, said second motion compensation prediction processing section being adapted to obtain an lower-layer motion compensation prediction signal upon reducing the number

of data;

a decision section for deciding in motion compensation on the basis of the lower motion compensation prediction signal whether motion compensation is correct;

a selector for selecting the upper-layer motion compensation prediction signal in response to a correct motion compensation prediction from said decision section, and the lower-layer motion compensation prediction signal in response to an incorrect motion compensation prediction; and

an encoder for encoding a motion compensation prediction value selected by said selector.

14. An apparatus according to claim 13, which includes means for detecting a motion vector for each object and the background in a block containing a plurality of objects or the background, and a motion vector detection range in the block containing the plurality of objects or the background is narrower than that in a block containing one object or the background.

15. An apparatus according to claim 13, further comprising motion vector detection means for detecting a motion vector in units of objects and performing motion vector search operation in the same object.

16. A video encoding apparatus comprising:

an orthogonal transform circuit for dividing

an input video signal into a plurality of blocks each containing N × N pixels and orthogonally transforming an arbitrary-shape picture in units of blocks to obtain a plurality of transform coefficients;

means for encoding and outputting an alpha-map signal for discriminating a background of a picture from an object thereof;

means for calculating an average value of pixel values of an object portion using the alpha-map signal in units of blocks;

means for assigning the average value to a background portion of the block;

means for deciding using the alpha-map signal whether a pixel in the object is close to the background;

means for compressing, about the average value, the pixel in the object decided to be close to the background; and

means for orthogonally transforming each block to output an orthogonal transform coefficient.

17. A video encoding/decoding system comprising:

a video encoding which includes,

an orthogonal transform circuit for dividing an input video signal into a plurality of blocks each consisting of N × N pixels and orthogonally transforming an arbitrary-shape picture in units of blocks to obtain a plurality of transform coefficients,

means for encoding and outputting an alpha-map signal for discriminating a background of a picture from an object thereof,

means for calculating an average value of pixel values of an object portion using the alpha-map signal in units of blocks,

means for assigning the average value to a background portion of the block,

means for deciding using the alpha-map signal whether a pixel in the object is close to the background,

means for compressing, about the average value, the pixel in the object decided to be close to the background, and

means for orthogonally transforming each block to output an orthogonal transform coefficient; and

a video decoding apparatus which includes,

means for decoding the alpha-map signal encoded by said video encoding apparatus,

means for decoding a block from the orthogonal transform coefficient,

means for calculating an average value of the object portion, the background portion, or the entire block,

means for deciding using the alpha-map signal whether the pixel in the object is close to the

background, and

means for decompressing, about the average value, the pixel in the object decided to be close to the background by said deciding means.

18. A video encoding apparatus comprising:

an orthogonal transform circuit for dividing an input video signal into a plurality of blocks each containing N × N pixels and orthogonally transforming an arbitrary-shape picture in units of blocks to obtain a plurality of transform coefficients;

means for encoding and outputting an alpha-map signal for discriminating a background of a picture from an object thereof;

means for calculating an average value of an object portion in units of blocks for the picture, using the alpha-map signal;

means assigning the average value obtained by said means for calculating the average value to a background portion of the block;

means for deciding using the alpha-map signal whether a pixel in the background is close to the object;

means for correcting the pixel in the background which is decided to be close to the object as a decision result of said deciding means in such a manner that a pixel value of the pixel comes close to a pixel value of the object near the pixel; and

means for outputting an orthogonal transform coefficient obtained by orthogonally transforming the block.